Taxonomy, Trees, and Truth in Historical Mammalogy

Robert J Asher
Department of Zoology, University of Cambridge

Address for correspondence:
Museum of Zoology, Downing St, Cambridge CB2 3EJ, United Kingdom
r-dot-asher-at-zoo.cam.ac.uk

Chapter for publication in Zachos F, Asher RJ. [anticipated 2017]. *Handbook of Zoology: Mammalia.* Walter deGruyter: Berlin.

Citation: Asher RJ. 2017. Taxonomy, Trees, and Truth in Historical Mammalogy. In Zachos F, and Asher RJ. Handbook of Zoology: Mammalia. Berlin: Walter deGruyter.

Abstract

For millennia, taxonomists have tried to make sense out of the world by visually and verbally organizing it. In biological terms, pre-evolutionary thinkers have often suggested an affinity of humans to other mammals, among other relationships expressed in their taxonomies. The availability of a well-corroborated tree for living mammals enables me in this chapter to ask if past authors succeeded (or not) in creating accurate taxonomies, and if specific methodological and theoretical advances (e.g., the recognition of evolution via descent with modification) are associated with improved taxonomies. They are. I find that different methods used over past centuries for building taxonomies (including pre-evolutionary, evolutionary, cladistic, and molecular) improves on its predecessor, although at varying rates and levels of significance. H.M.D. DeBlainville was an outlier among pre-evolutionary authors and produced substantially more accurate taxonomies than his contemporaries. Evolutionary authors constructed significantly more accurate taxonomies than their predecessors, and began at a higher level of accuracy than would be expected based on previously observed improvement over time. Cladistic methods show improvement at a greater rate than previous methods, but do not start at a level greater than expected based on improvement evident among pre-cladistic authors. Authors using molecular data, including datasets independent of the nuclear DNA key for recognizing the currently well-corroborated tree, began with levels of accuracy similar to those applying cladistic methods to morphology, but thereafter showed the greatest improvement in the shortest amount of time. All evolutionary methods (including cladistic

and molecular) significantly outperform pre-evolutionary methods in recognizing well-corroborated mammalian groups.

Introduction

Categorization is integral to rationality. A rational being observes its environment and tries to make sense of it: safe vs. dangerous; edible vs. indigestible; interesting vs. boring. These intuitive dichotomies are among the most basic forms of taxonomy, or systems for categorizing the patterns we observe. Taxonomies pervade human cultures, and are not only a result of real patterns but are also influenced by our social and psychological biases. Taxonomies in some fields are widely held and of great importance (e.g., law and economics), but are heavily influenced by arbitrary social convention, often to a greater extent than by any intrinsic reality to the things being categorized. In principle, biological patterns are amenable to categorization based on intrinsic reality. Anyone can perceive nature, and perceptions of biological patterns are often consistent across observers. Therefore, taxonomies that represent biological patterns can also be consistent, since they represent a reality that is accessible to independent observers, despite cultural and linguistic differences (Hunn 1975). When they do not represent reality, it may be because of limitations to symbolic language, economic or political motivations that are not biological, or misunderstandings of the process by which the phenomena to be categorized have become diverse.

We expect that when the process, or mechanism, behind a given phenomenon is well understood, taxonomies of that phenomenon will be more accurate than taxonomies based on a poorly understood process. This is not to imply that any particular theory of process needs to be assumed, endorsed, or even understood by those engaged in taxonomy, but recognizes a logical independence between efforts to categorize patterns, such as taxonomy, and the processes that may underlie those patterns, such as evolution (Patterson 1988). In those cases when a causal relationship really does exist between the two, understanding the process should make categorizations of its resulting patterns less susceptible to bias and more rooted in reality. For example, medical treatments based on the zodiac are less effective that those based on the germ theory because the latter entails not just claims of correlation or agency, but well-corroborated theories of cause and effect. Both astrology and the medical study of infectious disease entail descriptions of pattern (or "what"), but the latter goes further to explain the process (or "how"), and thereby leads to categories of treatment firmly rooted in empirical reality. An understanding of the process behind infection underlies a medical taxonomy including hygiene and antibiotics, whereas no demonstrable process is behind a medical taxonomy including Jupiter and the pituitary gland.

Similarly, the discovery of a process by which livings things have become diverse over time (i.e., evolution) should lead to an improvement in taxonomists' ability to categorize organisms. Animal

taxonomies are ubiquitous throughout recorded human history (Gregory 1910), whereas a widely accepted theory of biological evolution is fairly recent. This makes it possible to ask two interrelated questions: 1) are taxonomies based on an evolutionary understanding of nature more accurate than non-evolutionary taxonomies, and 2) has better access to biological patterns, and the methods to quantify them, led to improved taxonomic accuracy?

Goals of this study

In this chapter, I address these questions with a focus on mammals, and rely heavily on Gregory's (1910) history of mammalian classification, supplemented by a number of additional classifications and phylogenetic studies. If it is true that taxonomies of mammals have become better at representing reality since the widespread acceptance of evolution, and as access to pattern documenting that reality has improved, then we would expect two things to occur over time: 1) taxonomies should converge towards the currently most well-corroborated hypothesis of mammalian interrelationships, and 2) taxonomies published by authors using methods that provide a clearer grasp of reality should resemble each other more than those using inferior methods.

I argue that both expectations hold even though past taxonomies were generally not intended to represent evolutionary history. Al-Jahiz in the 9th century was an advocate of monotheistic creation (Eisenstein 1991; Stott 2012; Montgomery 2013). (Incidentally, contemporary religious creationists who offer explanations in the form of an agent, e.g., "god created species", often have views that are incompatible with a process-oriented explanation such as descent with modification [Numbers 2006]. However, without invoking any processes or mechanism by which their agent did the creating, and religious creationists do not provide much detail on this point, simple assertion of agency does not necessarily conflict with a theory of process such as evolution; see Sober 2008; Asher 2012.) Georges Cuvier was an empiricist and skeptical of Lamarckian evolution, but did not clearly advocate any specific theory of process either (Rudwick 1997). William King Gregory, in contrast, did accept evidence for evolutionary descent with modification as the process behind biodiversity (Gregory 1910). Despite this variation, al-Jahiz, Cuvier, and Gregory all used, at least in part, biological patterns intrinsic to the organisms they observed to construct their taxonomies. Whatever processes they thought were driving the patterns they observed, each author was constrained by the reality of those patterns. A bovid's split hoof was no less obvious to al-Jahiz than it was to Cuvier or Gregory. A process like evolution which proposes to explain differing levels of similarity across living things is, therefore, at least partly manifest in any taxonomy that uses similarity. We expect more recent taxonomies to be better than older ones for two reasons: First,

recent investigators benefit from more technology with which they can appreciate pattern; and second, recent investigators have a better understanding of the now well-corroborated process behind biological diversity (evolution).

Perhaps more interestingly, we do not have to assume that Gregory's understanding of an evolutionary process is accurate, but based on the patterns he interpreted to make claims about relationships, we can test whether it is or not (Penny et al. 1982; Patterson 1988). If he was right, then we would expect patterns of biological data unknown to Gregory (e.g., DNA) to yield taxonomies closer to his than to others who did not subscribe to an evolutionary process, such as Cuvier (1817) or Owen (1868). We would also expect that, if they are indeed better, novel methodologies such as cladistic taxonomy (Hennig 1966) would lead to improved accuracy and more agreement on classification among authors who use such methods.

In order to use the patterns evident in classifications to test a given process behind them, one cannot logically pre-suppose that process to have caused the pattern in the first place (Rieppel 1986; Patterson 1988). Hence, I use the terms like "cladogram", "branching diagram", "classification", and "taxonomy" to indicate past attempts at summarizing biological patterns, regardless of a given author's views on a theory of process. A "tree" or "phylogeny", on the other hand, denotes a pattern generated by an evolutionary mechanism (Patterson 1988). Therefore, comparisons of cladograms can test evolution; comparisons among trees cannot. This may seem like an esoteric point, and to some extent it is. Evolutionary theory has already been reasonably demonstrated; just as civil engineers may assume that gravity is relevant to their hypotheses on optimal bridge construction, contemporary biologists may assume that descent with modification is relevant to their hypotheses on how animals should be classified. On the other hand, terminological precision is important, and 21st century biologists should not forget the reasons why evolutionary descent with modification is a compelling theory of process. This theory entails postulates such as common descent that can be demonstrated by making predictions about branching diagrams (Penny et al. 1982) which do not have to presuppose anything about evolution (Patterson 1988). This makes the distinction between "cladogram" as a statement about pattern, and "tree" as a statement about process, worthwhile.

Methods

I extract cladograms from a number of historical classifications (Table 1; Appendix 1) and quantify the extent to which they resemble each other and the current, well-corroborated tree of mammals (Fig. 1). By "well corroborated", I mean a hypothesis of interrelationships among the ca. 5500

species of extant monotremes, marsupial, and placental mammals that has received support from a wide variety of methods and data (Asher et al. 2009). This tree is nearly but not completely resolved, and is based on consistency across topologies generated by alignments of 35,603bp (basepairs) of nuclear DNA for 169 mammals (Meredith et al. 2011), up to 15,535bp of nuclear and mitochondrial DNA for 1,265 rodents (Fabre et al. 2012), 43,616bp of nuclear and mitochondrial DNA for 193 marsupials and 10 other mammals (Mitchell et al. 2014), and 32,116,455bp of nuclear DNA for 36 mammals and 16,050bp of micro-RNA for 39 taxa (Tarver et al. 2016). The latter study also incorporated large, amino-acid sequence datasets from Hallström & Janke (2010), Romiguer et al. (2013), and O'Leary et al. (2013). Thus, Fig. 1 shows areas of agreement among the optimal topologies published by Meredith et al. (2011), Fabre et al. (2012), Mitchell et al. (2014), and Tarver et al. (2016). Polytomies are shown where these studies disagree (e.g., on sister taxa of Scandentia, Perissodactyla, and Chiroptera).

Quantifying similarity

I measured similarity by counting the number of "shared partitions" among rooted cladograms in Mesquite (Maddison & Maddison 2015), excluding the two basal-most nodes. This enables comparisons of cladograms that share similarly grouped taxa, but have different taxon samples. For example, as shown in Fig. 1, Storr (1780) places the otter (*Lutra*) close to other carnivorans, such as *Mustela*. This contrasts with authors such as Scopoli (1777) who placed *Lutra* with other aquatic species such as seals and hippos. A mustelid-*Lutra* group is therefore one area of agreement of Storr's classification with the well-corroborated phylogeny shown in Fig. 1. Similarly, John Ray (1693) accurately recognized humans with "simians" and bovids with cervids (Fig. 2). Slightly more complicated is Storr's group "Plantares" (in turn a part of his "Manuati" and "Primates"), consisting of *Didelphis* and *Phalanger*. I understand both terms to mean modern genera of marsupials, and Storr classifies no other marsupial taxon in his 1780 classification (Gregory 1910: 49-50). Thus, even though our modern understanding of marsupial interrelations does not have a *Didelphis-Phalanger* clade (Fig. 1), it does have a group containing the parent taxa of those two species (Didelphimorphia and Diprotodontia) to the exclusion of all other species classified by Storr. The "shared partitions" tool in Mesquite accurately represents this similarity, as depicted in Fig. 1, and I use it to compare each of the cladograms derived from historical classifications to the well-corroborated topology in Fig. 1, and also all pairwise comparisons across cladograms extracted from the 55 classifications listed in Table 1. All cladograms derived from these studies are available in nexus format in Appendix 1. For data comparisons and summaries I used R version 3.1.0 (R core team 2014) and Libreoffice/Microsoft Excel.

Comparisons among classifications need to account for differing taxon samples and the extent to which a given study made testable claims about affinity. For example, Prothero (2007) effectively sampled all of the 59 mammalian taxa represented here (as did McKenna 1975; see below), but only made precise statements of affinity for 37, leaving the rest in polytomies. In contrast, Song et al. (2012) sampled 27 of these taxa, resolved all of them, among which 20 groups are held in common with the well-corroborated topology (Fig. 1). In absolute terms, Prothero (2007) has more (24) partitions in common with Fig. 1, but as a proportion of the number he could have correctly identified given his sample (equivalent to the number of sampled taxa minus three), his accuracy is lower: 24 of a maximum possible of 56, or 0.43. Song et al. (2012) recovered 20 of a possible 24 in their sample, or 0.83. Hence, I evaluate historical topologies primarily using this ratio of correctly identified partitions by the number of potentially identified partitions for any given study.

In addition, many historical classifications did not precisely identify all of their terminal taxa, and I exercised discretion to give credit (or not) to authors who would have agreed with certain relationships given then-accepted knowledge of mammalian systematics, for example, that *Lutra* is a mustelid, *Melursus* an ursid, and *Hydrochoerus* a caviid. McKenna (1975) did not explicitly classify *Daubentonia* as a strepsirhine primate, or macropodids as diprotodont marsupials, but as a leading mammalogist of the latter half of the 20th century, he certainly knew about these taxa and would not have questioned these relationships. His 1975 classification is thus here represented as containing *Daubentonia* as part of Strepsirhini and Macropodidae as part of Diprotodontia. McKenna's 1975 classification did not explicitly mention *Dromiciops*, but he was of course aware of this South American marsupial as well. However, agreement did not yet exist in 1975 about the affinity of *Dromiciops* with Australian and not other South American marsupials; therefore, I credit his 1975 classification only with recognizing that *Dromiciops* is a marsupial, with uncertain affinities (in the form of a polytomy) to other marsupial groups. In other cases, past assumptions of monophyly among terminal taxa in older classifications were wrong, such as terrestrial Artiodactyla to the exclusion of cetaceans (Gatesy et al. 1996), "Insectivora" including tenrecids and chrysochlorids (Stanhope et al. 1998), or Microchiroptera including rhinolophoids (Teeling et al. 2005). Therefore, I represent the relevant terminals (e.g., Cetacea, Tenrecidae, Chrysochloridae, Rhinolophoidea) so as to identify those analyses that, even into the 21st century (e.g., Prothero 2007) misidentified several, now well-corroborated branches on the mammalian tree. In every case, my goal was to identify the biological groups stated or implied by a given author and make them comparable across cladograms inferred from historical classifications (Table 1).

Source classifications

Table 1 outlines the 55 publications that contain the taxonomies considered here, dating to Leviticus in (or near) the 6th century BC to Tarver et al. (2016). The individual classifications included here require a consistent and reasonably diverse number of animal groups in order to be compared across authors. For most pre-20th century classifications, I rely on Gregory (1910). For example, the summary and tables given for Ray (1693), both in his original publication (Fig. 2; Table 2) and in Gregory (1910: 18-19), are sufficiently detailed so as to compare Ray's ideas to his successors in later centuries. In contrast, Gregory's summaries for the Assyrians, Aristotle, Gessner, Caesalpinus, Wotton, Erxleben, Perrault, Buffon & Daubenton, Boddaert, Vic D'azir, Geoffroy St. Hilaire, and Duméril do not provide enough information to represent their ideas in a form that is sufficiently comparable to other studies; these authors are therefore excluded. In sum, I examine 55 branching diagrams that represent the patterns observed by 45 different authors (Appendix 1).

Most of the pre-1910 citations are derived from summaries thereof in Gregory (1910); the exceptions (e.g., Al-Jahiz [Eisenstein 1991]; Haeckel 1866, 1905; Gadow 1899) are given in Table 1. Several pre-modern and 20th century cladograms were also drawn from Pietsch (2012). Transforming older classifications into explicit branching diagrams is not a straightforward task; as noted above, they did not intend to formulate evolutionary trees. Ancient classifications such as those in Leviticus (ca. 6th century BC) or *The Book of Animals* by Al-Jahiz (9th century AD) were based on function and/or perceived utility for humans. Al-Jahiz used categories such as swimming, flying, and crawling to group, for example, bats with birds (Eisenstein 1991). The authors of Leviticus sought to distinguish clean vs. unclean animals for human consumption. Nor were later natural philosophers primarily concerned with representing patterns of common ancestry, which were either unknown (e.g., John Ray in the 17th century) maligned as unempirical (Cuvier reacting to Lamarck in the early 19th century), or otherwise disregarded until 1859. Even explicitly evolutionary taxonomies (e.g., Gregory 1910 or Simpson 1945) did not adhere to the modern convention of naming monophyletic taxa. Nonetheless, all of these classifications were based on at least some then-current ideas of animal similarity, whether in an evolutionary sense (e.g., Huxley 1872) or as empirical (Cuvier 1817) or philosophical (Owen 1868) attempts to document "natural" groups. Whatever the motivations behind any historical author may have been, and whatever their understandings of natural or supernatural processes were, there appears to have been (at least from the Renaissance onwards) a genuine desire to uncover "true" similarities across animal groups, and to use classification to reflect this similarity. This makes pre-modern classifications fair game to

investigate how the methods and theories of their day compare to those of the present.

Interpreting historical classifications

Early taxonomists sometimes used terms that imply different animals relative to today's standard. For example, Lacépède (1799) used "*Cynocephalus*" to refer to baboons (Primates), and "*Galeopithecus*" for colugos (Dermoptera)."Manati" or "Manatus" were probably used by Ray (1693), Klein (1751), and Scopoli (1777) in reference to pinnipeds, not sirenians, as these authors classified "Manati" using features of the pes (completely lacking in Recent sirenians). For Ray, "Manati" was part of his "Quadrupedia" (Fig. 2; Table 2), and Klein referred to its hindfeet (along with those of the otter, beaver, seal, and walrus) as duck-like: "pedibus quibuscumque anserinis", presumably in reference to the hindfoot webbing of aquatic carnivorans and rodents that would have been obvious in the descriptions and/or cadavers available to 17th and 18th century observers. However, Pennant (1781) and Bonaparte (1837) probably used "*Manati*" for Sirenia, as they describe them as "herbivorous" and "phytophagous" (respectively). To make matters yet more interesting, authors such as Lacépède (1799) used "*Manatus*" in reference to Atlantic sirenians (i.e., manatees), but understood their current genus name (*Trichechus*) to be a walrus ("*Trichechus" rosmarus* instead of *Odobenus rosmarus*). Cuvier (1800) and Illiger (1811) also identified "*Manatus*" among manatees ("lamatins") and "*Trichechus"* among walruses ("morses").

More generally, we take for granted the affinity of the otter with mustelid carnivorans, of kangaroos with diprotodont marsupials, of the capybara with caviomorph rodents, and of the aye-aye and human with other primates. In contrast, early taxonomists could not. In the seventeenth century, John Ray would never have heard of, much less seen or dissected, an elephant shrew, golden mole, colugo, numbat, or platypus. The novelty of these animals is sometimes reflected in early classifications, such as Lacépède's (1799) placement of the hyrax among rodents and the aye-aye with kangaroos, or Illiger's (1811) placement of "thumbed" marsupials with primates in his "Pollicata".

Yet the anatomical patterns which form the basis of most historical classification were sufficient to enable some investigators to recognize startlingly accurate groups, even in the absence of a theory of process as to how those patterns arose. Ray (1693) classified humans and at least some apes as many-toed unguiculate animals with flat-nails together in his category "simians" (Fig. 2; Table 2). Subsequent authors, even into the 20th century, tended to place humanity in its own category apart from other animals (e.g., the Kingdom "Psychozoa" of Huxley 1957:91). Cuvier (1800) was among

the first to recognize the affinities of the hyrax near elephants (albeit with a variety of other "ungulates" mixed in), and Illiger (1811) and DeBlainville (1834) placed the aye-aye with other primates. To give due credit to these prescient categorizations, I have represented the taxa most frequently misplaced by early authors (e.g., *Daubentonia, Lutra,* Macropodidae) as entities apart from the high-level taxa to which they belong (Strepsirhini, Musteloidea, Diprotodontia). This enables me to graphically acknowledge taxonomies in which these species are misplaced (e.g., Illiger's "Pollicata"), and give due credit to those which are correct. I do not thereby intend to imply that the Strepsirhini of, say, Meredith et al. (2011) is paraphyletic by showing it apart from *Daubentonia* (Fig. 1).

As graphic depictions of evolutionary lineages became more common in the 19th and early 20th centuries (Pietsch 2012), authors did not necessarily make their branching diagrams consistent with their classifications. For example, Gregory (1910) illustrated his preferred branching diagram in figures 31 and 32 of his 1910 monograph. By contemporary standards, his depiction of "diprotodont" and "polyprotodont" marsupials in his fig. 32 implies marsupial paraphyly, as the latter branch closer to placental mammals than the former. Yet his fig. 31 shows the two immediately adjacent, reflected also in his classification (1910: 464). Based on his classification and fig. 31, I have represented Gregory's view as supporting marsupial monophyly. Similarly, Gregory uses the separate taxa Pinnipedia and Fissipedia to delineate aquatic vs. terrestrial carnivorans, but his figure of their "genetic relations" (his fig. 31) shows a more modern placement of seals and walruses (i.e., pinnipeds) among ursoids and other caniforms to the exclusion of feliforms. Hence, the cladogram representing his views (Appendix 1) reflects this evolutionary understanding more than his taxonomy, which of course predated the widespread acceptance of cladistic taxonomy and the use of monophyletic taxa. In addition, Gregory's classification (1910: 464-466) delineates groups such as "Therictoidea", containing his "Insectivora" plus Carnivora. Yet in his depiction of the "genetic relations of the orders" (1910: fig. 31), he places lipotyphlan families on a branch adjacent to Archonta. Again, I rely on the latter to represent his views, and where possible, I use authors' own figures (as indicated in Table 1) to represent their views here.

Winge (1941) is the posthumous, English translation of several papers originally published in Danish between 1887 and 1919, apparently edited by the author as late as 1921 (1941:vii), shortly before his death in 1923. The 1941 volume therefore substantially post-dates his actual taxonomic views, and for comparative purposes I date his work at 1921 (Table 1). Winge's volume outlined mammalian comparative anatomy and systematics, focusing on monotremes, marsupials, bats,

"insectivorans", and "edentates". His ideas on the interrelations among "edentates" are figured on p. 329, bats on p. 261, "insectivorans" (including scandentians and dermopterans) on p. 145, marsupials on p. 70, therians vs. monotremes on pp. 16-17, and he briefly shows his understanding of "Ungulata" in relation to other orders on p. 136. Winge did not detail his views on the classification of the remaining mammalian groups beyond the single sketch on p. 136. Thus, in order to represent his ideas and compare them to other classifications, it is necessary to represent Glires, Primates, Carnivorans, and "ungulates" as polytomies. This places his classification at a disadvantage compared to others from the early 20th century, as the number of potentially recognized clades in common with the well-corroborated tree is high, but he does not give details on what he thought the relations were within these four groups, thereby missing potentially correct groups recognized by others in the late 19th and early 20th century (e.g., pinnipeds, ruminants, perissodactyls). Nonetheless, his classification improves the sample of those from the early 20th century, so I include his work in Table 1.

Unlike Gregory (1910), Simpson (1945) did not present a figure of his views on mammalian phylogeny. Despite the many qualifications in his text, and the fact that as an "evolutionary taxonomist" (Van Valen 1978) Simpson did not follow the standard of monophyletic taxa generally practiced today, his classification (pp. 39-162) is a definitive and unambiguous statement based largely on his understanding of mammalian phylogenetic history. The cladogram representing his views is therefore based on his nested classification, not on the many more nuanced statements about relationships in his text. In his classification, for example, he places tupaiids and tarsiiforms within "Prosimii", and aquatic carnivorans in "Pinnipedia" to the exclusion of terrestrial carnivorans. Both concepts were widely held in the middle 20th century, although Simpson was often more pragmatic than dogmatic about the biological reality of any given taxon.

In *Vertebrate Paleontology* (1st edition 1945, 3rd edition 1966), Alfred Sherwood Romer figured his ideas on mammalian systematics. These were re-published in Pietsch (2012: 190-194) for Placentalia, Carnivora, and Artiodactyla (1945) and Rodentia (1966). The classification from Romer (1959) is derived from *The Vertebrate Story* (fourth edition), using his figures for synapsids (p. 231), placentals (p. 237), carnivorans (p. 240), perissodactyls (p. 264), artiodactyls (p. 265), and primates (p. 310).

The final volume (#17) of the comprehensive *Traité de Zoologie* (Grassé 1955) consisted of two sub-volumes (or "fascicules") dedicated to systematics and comparative anatomy of mammals.

Pierre Paul Grassé (1895-1985) was editor of this massive series; its 17 volumes contained from one to seven fascicules and were written by a number of additional authors, published between 1952 and 2007. Grassé was known for his sympathy for Lamarckism, as described in his 1973 book *L'évolution du vivant* (published in English as *Evolution of Living Organisms*). The first figure in the opening chapter of *Traité de Zoologie* vol. 17 fasc. 1 is an outline of what was then understood about the relationships among the extant eutherian orders, but was composed by the evolutionary paleontologist Jean Piveteau, not Grassé. This figure forms the basis of the cladogram extracted here. Further details on the interrelationships within specific orders follow the classifications given in the subsections therein, such as "ungulates", carnivorans, and whales in part I and "edentates", "insectivorans", primates, glires, and bats in part II.

Novacek (1986) published a classification and figured four different cladograms based on varying combinations of his cranial dataset, and used parsimony as his optimality criterion to figure those cladograms. His fig. 35 is based on the most characters and is used here to represent his views. Prothero (2007) is a popular exposition of paleontological evidence that demonstrates evolution. In his book, he figures a cladogram of placental orders (p. 285), and others for carnivorans (p. 291), primates (p. 338), artiodactyls (p. 312), tethytheres (p. 324), perissodactyls (p. 309), and "ungulates" (p. 298). Where these differ (e.g., sirenians and hyracoids) I follow his figure 13.9 (p. 285).

Defining methodologies

I categorize the studies listed in Table 1 into one of seven types: ancient, pre-evolutionary, evolutionary, cladistic, molecular, combined, and correct. Two molecular studies used to define the well-corroborated tree (Meredith et al. 2011 and Tarver et al. 2016) are defined as "correct", apart from "molecular", in order to measure the latter category without circularity. The two oldest classifications (Leviticus and Al-Jahiz) are defined as "ancient", not "pre-evolutionary", as they were based more on perceptions of function and human utility than other factors (e.g., anatomy). All of these categories are practical groupings, and I acknowledge that they are not mutually exclusive and that authors in each were not singular practitioners of their respective methodology. For example, McKenna (1975) applied "cladistic" principles to nomenclature, but built his classifications in a similar manner as "evolutionary" authors such as Simpson (1945).

Authors in this sample from the Renaissance to the mid-19th century used "pre-evolutionary" criteria to build their classifications. While such criteria are diverse, and in later cases (e.g.,

DeBlainville 1834; Owen 1868) were probably influenced by proto-evolutionary ideas, even the oldest pre-evolutionary study (Ray 1693) consisted of more than just dietary prescriptions and superficial functional categories. Ray (1693) recognized that mammals, including humans and whales, breathed air, possessed two ventricles in the heart, and had live birth ("*pulmone respirantia, corde ventriculis praedito duobus, vivipara*"; Fig. 2), demonstrating that Ray valued comparative anatomy as a basis for classification. Authors (except Owen) from 1859 through Romer (1959) used "evolutionary" criteria, i.e., they understood their taxonomic groups to reflect snapshots of evolving mammalian lineages. While some variation exists among these authors (e.g., Grassé as noted above), those behind the classifications used here (Table 1) largely or entirely endorsed biological evolution in the form of the Modern Synthesis (Simpson 1944). Importantly, further elaborations and improvements upon it have never disproven its basic mechanism of descent with modification (Wray et al. 2014). Moreover, authors in the cladistic, combined, and molecular categories are also "evolutionary" in the broad sense, but incorporated additional methodological improvements, the utility of which is worth testing here.

McKenna (1975) is widely credited (e.g., Prothero 2007) as the first author to apply cladistic methods to a classification of mammals. In fact, he built his classification in a similar fashion as previous authors, using experience and intuition without applying an optimality criterion to discrete characters. However, unlike Romer or Simpson, McKenna used the now widely accepted standard of naming monophyletic taxa; his classification therefore qualifies as "cladistic". McKenna & Bell (1997) and Prothero (2007) followed suit, whereas other cladistic studies (Novacek 1986, 1992; Shoshani & McKenna 1998) favored cladograms that minimized character change, defined by a matrix of discrete morphological characters. My 2003 study (Asher et al. 2003) is also cladistic in both its taxonomy and phylogeny-reconstruction method, but combines a 22kb molecular dataset with 196 morphological characters, and so is categorized as a "combined" analysis. Studies in the "molecular" category also name only monophyletic taxa, but differ in that their favored hypotheses are based on DNA characters alone, evaluated primarily with probabilistic optimality criteria (e.g., maximum likelihood, Bayes' theorum) using models of sequence evolution (Felsenstein 2004).

Results

The main question posed in the introduction was whether or not taxonomies of mammals have become biologically more accurate with an evolutionary understanding of biological process, and with improvements in technologies and methods to recognize patterns. If they have, we would expect that 1) taxonomies should converge towards the currently most well-corroborated hypothesis

of mammalian interrelationships, and 2) taxonomies published by authors using improved methods should become more similar to each other over time, as we would expect improved methods to enable taxonomies to more clearly reflect biological reality. By and large, the historical classifications quantified here bear out both expectations. The two periods of time that show the greatest increases in both similarity of historical classifications to the currently best corroborated tree (Fig. 3), and in the agreement of authors with each other (Fig. 4) are the mid 1800s and the late 1990s. This corresponds to the increased acceptance of evolution post-1859 and the widespread use of molecular data in the 1990s.

Similarity to the known tree over time

A best fit line to data points falling into each of the methodological categories with adequate samples (pre-evolutionary, evolutionary, cladistic, and molecular) shows that there is increased accuracy over time in each (Figs. 3, 5), although not necessarily with levels of significance above 0.95 (Fig. 3). Nonetheless, the pre-evolutionary trendline is below the accuracy recovered in the four earliest evolutionary studies (Haeckel 1866, Gill 1870, Huxley 1872, and Flower 1883), suggesting that early evolutionary studies performed better than expected based simply on improvement over time. The four least accurate evolutionary studies (Cope 1898, Haeckel 1905; Osborn 1917; Winge 1921) are close to the pre-evolutionary trend line (Fig. 3) extended into the early 20th century. Trendlines for evolutionary and cladistic studies intersect the trendlines of their successors (cladistic and molecular, respectively). This suggests two interpretations: first, that evolutionary studies were fundamentally better at recovering a biologically real mammalian taxonomy than pre-evolutionary studies (as expected if evolution actually happened), and second, that at least some of the improvements from cladistic and molecular methods might have been accessible to those using just (respectively) evolutionary and cladistic methods, at least initially. However, cladistic and molecular methods exhibit steeper trendlines (Fig. 3) showing that the rate at which improvements happened using these latter methods, particularly molecular, was faster than rates in pre-evolutionary and evolutionary methods.

Of the 20 classifications in this study that pre-date 1859, only DeBlainville (1816, 1834) approaches evolutionary classifications in the number of accurately/potentially reconstructed groups. He was the first to recognize that monotremes, marsupials, and placentals are distinct; he also recognized the integrity of glires and primates, and that suiforms and ruminants are closer to each other than to, say, perissodactyls. DeBlainville's 1834 classification (Figs. 6b, 7) was among the first to recognize seals and walruses (pinnipeds) in a group near felids and canids. His ratios of correct/potential

groups (1816: 0.29, 1834: 0.28) are better than several derived from classifications published decades later, including those by Huxley (1872, 0.24), Flower (1883, 0.27), Cope (1898, 0.19), Haeckel (1905, 0.21), Osborn (1917, 0.23), and Winge (1921, 0.18). Otherwise, the most accurate pre-evolutionary classification is that of Owen (1868, 0.22), followed by John Ray (1693, 0.14).

The most accurate evolutionary classification is Gregory (1910), who identified 23 groups in common out of a maximum possible of 56 (0.41, Fig. 6c). This included strepsirhine and haplorhine primates together to the exclusion of scandentians, pinnipeds closer to caniform than feliform carnivorans, hyracoids near proboscideans and sirenians, and rodents and lagomorphs together in Glires. The most accurate cladistic classification is McKenna & Bell (1997), who recovered 27 out of 56 possible groups (Fig. 6d). They improved on Gregory (1910) by placing tarsiers among haplorhine primates, pholidotans near carnivorans, dermopterans closer to primates than bats, the common ancestor of xenarthrans close to the placental root, and *Dromiciops* closer to Australian marsupials than to other South American taxa.

Inter-author agreement and methodological improvements

In order to measure the extent to which authors publishing in a given period of time agreed with one another (as would be expected if they were using effective methods and/or an improved understanding of process), I created 12 time bins with 4-5 classifications each, from 500BC to 2016 (Table 3, Fig. 4). Fig. 4 shows the extent to which the authors within each of these time bins agreed with one another by showing the median value of all pairwise combinations of groups held in common within each time bin. For example, from 1866 to 1883, I sampled five classifications (Haeckel, Owen, Gill, Huxley, and Flower). There are 10 pairwise combinations of these studies, ranging from 12 (e.g., Huxley-Flower) to 16 (e.g., Gill-Owen) groups in common with each other. The median of all 10 comparisons is 13 groups in common. When standardized by the ratio of observed to potential groups in common (see Methods), the median is 0.343 groups in common. This is substantially higher than the previous bin (0.129 in 1817-1855), and slightly higher than the following (0.276 in 1898-1910).

The two greatest increases in the median number of groups held in common among authors within each time bin took place in the mid-19th and near the beginning of the 21st century (Fig. 4). Two decreases in inter-author agreement also occur: from 1883 into the 20th century and from 1997 into the 21st. These decreases might be explained as follows: Towards the end of the 19th century and into the 20th, authors recognized the need to categorize then-novel and obscure mammalian taxa

(e.g., *Solenodon*, *Dromiciops*, *Notoryctes*, macroscelidids, tupaiids), but were not sure how to do so. Hence, the number of potentially recognized groups in common with the well-corroborated tree rose in the form of polytomies, reducing the number of resolved groups in common with their peers. Stated differently, resolution generally dropped and polytomies increased from the 1860s into the 1900s. An explanation of the drop in inter-author agreement during the 1990s-early 2000s is that several molecular studies from the 1990s favored a root position for Placentalia within Glires (e.g., Stanhope et al. 1998) or erinaceids (Arnason et al. 2002). This led to a marked reduction in agreement with other branching diagrams (e.g., Murphy et al. 2001) that favored a root position closer to the common ancestors of Afrotheria (elephant, sea cow, hyrax, tenrec, golden mole, sengi, aardvark) and Xenarthra (sloth, anteater, armadillo) which resembles the correct root position near the clade uniting these two groups in Atlantogenata (Tarver et al. 2016; see Fig. 1).

During the time bin (1986-1997) when cladistic methods in nomenclature and morphology dominated, the authors sampled here did not show a greatly increased level of agreement compared to the previous bin (1945-1975; see Fig. 4), implying that cladistic methods were not a great improvement over evolutionary ones in terms of enabling investigators to create accurate taxonomies. However, although slight, similarity of cladistic authors' taxonomies to the well-corroborated tree does increase (Fig. 5; Table 4). Moreover, some time bins contain multiple methods. For example, Miyamoto & Goodman (1986) is a molecular study in a "cladistic" age; Prothero (2007) is a cladistic study in a "molecular" age. The (admittedly small) sample of authors that applied cladistic nomenclature and/or phylogeny reconstruction methods to morphological data (e.g., Shoshani & McKenna 1998) exhibited, on average, higher levels of similarity to the known tree compared to evolutionary authors (e.g., Haeckel 1866; Gadow 1898; Simpson 1931, 1945). Evolutionary authors were in turn more accurate, on average, than pre-evolutionary studies (e.g., Owen 1868; see Figs. 3 and 5). Ignoring for a moment the unavoidable non-independence across authors in Table 1 (see Discussion), and using a non-parametric Wilcoxon rank sum test (excluding the studies in Table 1 used to define the well-corroborated tree; Meredith et al. 2011 and Tarver et al. 2016), the better performance of evolutionary, cladistic, and molecular methods over pre-evolutionary studies is highly significant ($p<<0.01$); the higher accuracy of cladistic over evolutionary studies is suggestive but not significant ($p=0.068$); the higher accuracy of molecular over evolutionary studies is highly significant ($p=0.002$); and the higher accuracy of molecular over cladistic studies is suggestive but not significant ($p=0.100$; see Fig. 5; Table 4).

Molecular studies show considerable overlap with the well-corroborated tree, which is of course

based on the largest molecular datasets (see Introduction). The least accurate molecular study is also the oldest in this sample: the analysis of protein sequences by Miyamoto & Goodman (1986), which recovered 13 out of 39 possible (for their sample) well-corroborated groups, yielding a ratio of 0.36. They are followed by the 3-gene (12S, 16S, valine) analysis of mitochondrial RNA of Stanhope et al. (1998, 0.48), a combined DNA-morphology study of mine (Asher et al. 2003, 0.55), and the mitogenomic analyses of Arnason et al. (2002, 0.56; 2008, 0.59) and Kjer and Honeycutt (2007, 0.6). The branching diagram of McCormack et al. (2012) is among the most accurate, recovering 11 out of a possible total of 15 partitions (or 0.73). Notably, all of the preceding nucleotide analyses (Miyamoto & Goodman 1986 concern protein sequences) exceed the accuracy of those based on cladistic morphology methods. (Stanhope et al. 1998 has a ratio of shared/potentially shared nodes of 0.48387; McKenna & Bell 1997 has 0.48214). Studies by Stanhope et al. (1998), Arnason (2002, 2008), and Kjer and Honeycutt (2007) are based entirely on mitochondrial data; the dataset of McCormack et al. (2012) is comprised of ultraconserved elements (UCEs). These data did not contribute to the two studies of greatest relevance in defining the well-corroborated tree (Meredith et al. 2011 and Tarver et al. 2016; see Fig. 1). Song et al. (2012) used genomic data that do overlap with the Tarver et al. (2016) dataset, but employed different phylogenetic methods (coalescence) to identify their optimal branching diagram, which was very similar to the well-corroborated tree (0.83). Because mitochondrial data were used to help define the correct tree for marsupials and rodents by (respectively) Mitchell et al. (2014) and Fabre et al. (2012), and because some investigators (e.g., Song et al. 2012) have used genomic datasets that overlap those employed to identify the topology in Fig. 1 (Meredith et al. 2011; Tarver et al. 2016), I am compelled to avoid the circular conclusion that accuracy among molecular phylogenies improves over time, when that accuracy is in turn defined by some of those same molecular data. Yet the independence of mitochondrial data (e.g., Kjer & Honeycutt 2007) and UCEs (McCormack et al. 2012) from the largely nuclear coding regions behind the well-corroborated tree (Fig. 1) enables an only slightly less-strong version of this conclusion, as these independent, molecular studies are clearly more successful in recovering the well-corroborated tree compared to all of the branching diagrams implied or stated by authors prior to 2000, regardless of data or method (Fig. 3).

Discussion

Assuming that the current, well-corroborated hypothesis of mammalian interrelationships (Fig. 1) is in fact "the true tree", or reasonably close to it, then there is no question that human taxonomies of mammalian diversity have improved greatly since the authors of Leviticus. Long before the advent of PCR and methods to reconstruct evolutionary history, and even before the widespread acceptance

of evolution itself, human classifications of mammalian diversity show improvement (Fig. 3). Furthermore, there are correlations in the timing of this improvement with the development of specific advances in biology, particularly the first, clear articulation of the evolutionary mechanism in the mid-19th century, and the widespread application of molecular data to systematics in the late 20th century.

The improvement over time evident among evolutionary studies yields a trend line that fits squarely among cladistic studies in the 1980s, suggesting that other aspects of improved technology and access to pattern may have led to the better accuracy among studies in the 1980s, rather than any intrinsic advantages of cladistic methods. Relatedly, all authors considered here enjoyed insights and methods not considered in this study, and their appreciation of such here-unquantified factors undoubtedly played some role in the observed improvement over time (Fig. 3). Notably, at least some of the early 20th-century authors behind the classifications shown in Table 1 were sympathetic to Lamarckian mechanisms (e.g., Grassé 1955 and possibly Winge 1941). Osborn (1917) was notoriously unclear about his preferred evolutionary mechanisms. (However, according to Pietsch [2012], the figure in Osborn's 1917 volume was by Gregory.) The 19th century witnessed greatly improved anatomical knowledge, such as the dissemination of microscopic techniques that facilitated discoveries about previously inaccessible areas such as histology and embryology. All of this would potentially have helped taxonomists reach better conclusions about "natural" groups during the course of the 19th century. Nonetheless, as summarized above, correlations of improvement over time for a given method (Figs. 3, 5) show that the advent of evolutionary studies post-1859 were above where accuracy was predicted to have been given the trend evident among pre-evolutionary studies. Notably, the pre-evolutionary taxonomist Richard Owen (1868) is also above the trend line derived using other pre-evolutionary studies, but still below that derived from evolutionary studies (Fig. 3).

H.M.D. DeBlainville

The accuracy of Henri Marie Ducrotay DeBlainville (1816, 1834; Figs. 6b, 7) was even more impressive; his 1834 classification recognized nearly twice as many accurate clades compared to Illiger (1811), and nearly three times as many as Cuvier (1817; see Fig. 6a), despite the fact that both Illiger and Cuvier classified more taxa than DeBlainville. Gregory (1910: 81, 82) described DeBlainville's 1834 work as "brilliant, prophetic" and "far superior to any hitherto met with". As noted above, DeBlainville was more accurate than several evolutionary authors, and his ratio of actual to potentially recognized groups is even above an evolutionary trendline extended backwards

into the 1830s (Fig. 3).

DeBlainville (1777-1850) was about eight years younger than Georges Cuvier (1769-1832). His personality and relationship to Cuvier may shed light on why he stands out from others of his time. First, although he was initially a student of Cuvier, and spent his career in Paris, DeBlainville was too independent-minded to serve as a technician, and eventually reached his intellectual independence with some animosity towards his former mentor (Appel 1980). DeBlainville was also sympathetic to the "philosophical zoology" of Lamarck and Geoffroy St. Hilaire, who endorsed idealistic theories which Cuvier famously held in low regard (Appel 1987). Thus, DeBlainville was keen to set himself apart from Cuvier, who was notoriously anti-evolutionary, or at least anti-Lamarckian (Rudwick 1997).

Moreover, Georges Cuvier was arguably the most politically influential biologist of all time, and would have likely exercised substantial influence (deliberately or not) towards contemporaries who lacked DeBlainville's animosity. Other late 18th and early 19th century taxonomists such as Lacépède (1799), Illiger (1811), and Blumenbach (1830) agreed with Cuvier (1800, 1817) in not recognizing marsupials, monotremes, or placentals as distinctive groups, and in emphasizing hand and foot structure in classification, thereby placing such taxa as didelphid marsupials near primates and/or carnivorans (Fig. 6a). Thus Blumenbach's "Multungulata" (repeated by Illiger [1811]) resembled Cuvier's "pachydermes ordinaires" in joining tapirs, suiforms, and rhinoceratids with proboscideans and/or hyraxes. These authors also placed one or more aquatic mammalian groups together, variably including sirenians, cetaceans, pinnipeds, ornithorhynchids, and/or castorids. Bonaparte (1837) recognized placental mammals as distinct from monotremes and marsupials, but followed Cuvier in using a "pachyderm"-like group, and in placing sirenians with cetaceans and the aye-aye (*Daubentonia*) with rodents (Appendix 1). DeBlainville (1816, 1834) made none of these mistakes. It remains for future studies to decipher if this was because of his antipathy towards his senior rival, his affinity for Lamarckian ideas that Cuvier rejected, and/or because of his other insights into comparative anatomy.

Non-independence across studies

In addition to the observation above that some of the molecular studies share data with those used to define the well-corroborated tree, there is further interdependence among the studies considered here. The authors listed in Table 1 all sought to categorize nature, and any later author would have undoubtedly been influenced by those who came before. DeBlainville was a successor of Cuvier,

and would also have read Brisson, Linnaeus, Ray, Leviticus, and possibly even Al-Jahiz. All but the authors of Leviticus (ca. 6th century BC) would have potentially been influenced by Aristotle (4th century BC), and later authors will always have had more predecessors by whom to be influenced. In a similar vein, later authors in the modern era would have been more likely to use overlapping datasets, i.e., the same or similar character complexes or genes, to generate their branching diagrams. Indeed, this dataset (Table 1) contains ten pairs of branching diagrams derived from the same first author (Arnason, Blumenbach, Cuvier, DeBlainville, Haeckel, Linnaeus, McKenna, Novacek, Romer, and Simpson). Furthermore, as previously noted, the figure in Osborn (1917) was in fact drawn by Gregory (Pietsch 2012), a long-time protégée of Osborn at the American Museum of Natural History.

Overall, the median number of well-corroborated trees agreed upon by different authors increases over time. However, as noted above, there are dips from 1883 into the 20th century, and from 1997 into the 21st (Fig. 4), showing that authors are not simply repeating each other over time. Furthermore, regardless of how authors are influenced by their predecessors in deciding upon their favored classifications, similarity to the now well-corroborated mammalian tree (Fig. 1), derived from data that played no role in creating any of the branching diagrams in Table 1 prior to Miyamoto & Goodman (1986), is a discrete, independent metric by which to judge accuracy. From a purely chronological view, there are spans of time (e.g., 1693-1800 or 1920-1992) when similarity to the well-corroborated tree does not trend upward (Fig. 3). Nonetheless, the average success of classifications based on pre-evolutionary criteria was lower than those based on evolution, which was in turn lower than those based on cladistic morphology, which was in turn lower than those based on molecular data (Fig. 5; Table 4). This is consistent with the hypothesis that widespread acceptance of the evolutionary process following 1859 led to a genuine improvement in our ability to categorize mammals, as did the increased availability of molecular data starting in the 1990s. A smaller improvement in the ability of classifications to recover the known tree took place with the adoption of cladistic methods in nomenclature, although (as noted above) it is not clear if this improvement is due to cladistics itself or to other advances that took place during the mid-20th century, not directly accounted for in this study.

Future directions and conclusions

The sample considered here could undoubtedly be improved, for example by including more than six cladistic morphology studies and more than one pre-1990 molecular study. In addition, I cannot rule out the possibility that authors such as McKenna & Bell (1997) were influenced by the then

extensive molecular literature, for example on comparisons of protein sequences (e.g., Miyamoto & Goodman 1986) and/or DNA hybridization (Kirsch et al. 1991). It is difficult to quantify the extent to which McKenna's relatively accurate classification (by 1997 standards) was the result of molecular datasets "leaking" into it. McKenna was himself a student of molecular phylogenetics (McKenna 1991), but his classification was more accurate than (and far from identical to) the protein-sequence branching diagrams published by (for example) Miyamoto & Goodman (1986). McKenna had no reservations about naming groups unsupported by molecular studies (cf. Glires with macroscelidids or hyracoids close to perissodactyls), and he rejected claims made by some early molecular analyses (e.g., paraphyletic Rodentia and Odontoceti). Hence, neither he nor other cladistic authors were simply repeating results from molecular studies in their classifications.

In summary, the success of all evolutionary methods in recovering much of the well-corroborated mammalian tree are significant relative to pre-evolutionary methods (Fig. 5; Table 4). This is consistent with the hypothesis that the widespread acceptance of evolution in the mid-19th century, and application of further evolutionary methods and data in the 20th, helped mammalian taxonomies to accurately reflect reality, as would be expected if evolution actually happened.

Acknowledgments

I am grateful to Roberto Inchingolo and Adrian Friday for access to literature. I thank Adrian Friday, Nick Matzke, Marcelo Sánchez-Villagra, Caroline Willich, and Frank Zachos for helpful comments on the text, and Frank Zachos for his encouragement and patience with this chapter.

Literature Cited

Appel TA. 1980. Henri de Blainville and the animal series: A nineteenth-century chain of being. J. Hist. Biol. 13:291–319.

Appel TA. 1987. The Cuvier-Geoffroy Debate: French Biology in the Decades Before Darwin. Oxford University Press.

Arnason U, Adegoke J, Bodin K, Born EW, Esa YB, Gullberg A, Nilsson M, Short RV, Xu X, Janke A. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. Proc Natl Acad Sci USA 99: 8151–8156.

Arnason U, Adegoke JA, Gullberg A, Harley EH, Janke A, Kullberg M. 2008. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. Gene 421: 37–51.

Asher RJ, Bennett N, Lehmann T. 2009. The new framework for understanding placental mammal evolution. BioEssays 31: 853–864.

Asher RJ, Helgen KM. 2010. Nomenclature and placental mammal phylogeny. BMC Evolutionary Biology 10: 102.

Asher RJ, Novacek MJ, Geisler JH. 2003. Relationships of endemic African mammals and their fossil relatives based on morphological and molecular evidence. J Mamm Evol 10: 131–164.

Asher RJ. 2012. Evolution and Belief: Confessions of a Religious Paleontologist. Cambridge: Cambridge Univ Press.

Blumenbach JF. 1779. Handbuch der Naturgeschichte, mit Kupfern. Göttingen: Dieterich.

Bonaparte CL. 1840. A New Systematic Arrangement of Vertebrated Animals. Trans. Linn. Soc. London. 18:247–304.

Brisson MJ. 1762. Regnum animale in classes IX distributum, sive synopsis methodica sistens classium, quadripedum scilicet & cetaceorum, particularum divisionem in ordines, sectiones, genera & species. Leiden: Theodorum Haak.

Cabrera A. 1922. Manual de mastozoología. Madrid: Calpe.

Cope ED. 1898. Syllabus on the Vertebrata. Philadelphia: Univ Pennsylvania Press.

Cuvier G. 1817. Le règne animale, 4 vols. Paris: Déterville.

Cuvier G. 1800. Leçons d'anatomie comparée, 5 vols. Paris: Baudouin.

Eisenstein H. 1991. Einführung in die arabische Zoographie: Das tierkundliche Wissen in der arabisch-islamischen Literatur. Berlin: D. Reimer.

Fabre P-H, Hautier L, Dimitrov D, Douzery P, Emmanuel J. 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. BMC Evol Biol 12: 88.

Felsenstein J. 2004. Inferring phylogenies. Sunderland MA: Sinauer Associates.

Flower WH. 1883. On the arrangement of the orders and families of existing Mammalia. Proceedings of the Zoological Society of London 1883: 178–186.

Gadow H. 1898. A Classification of Vertebrata, Recent and Extinct. London: Black.

Gatesy J, Hayashi C, Cronin MA, Arctander P. 1996. Evidence from milk casein genes that cetaceans are close relatives of hippopotamid artiodactyls. Mol Biol Evol 13: 954–963.

Gill T. 1870. On the relations of the orders of mammals. In Proc. Amer. Assoc. Adv. Sci., 19th meeting, pp. 267–270.

Grassé P-P. 1973. L'évolution du vivant. Paris: Albin Michel.

Grassé P-P. 1955. Mammifères: les ordres: anatomie, éthologie, systématique. Tome XVII, Fascicule 1&2. Paris: Masson.

Gregory WK. 1910. The orders of mammals. Bulletin of the American Museum of Natural History 27:1–524.

Guérin-Méneville F-E. 1836. Dictionnaire pittoresque d'histoire naturelle et des phénomènes de la nature. Au Bureau de souscription: Paris.

Haeckel E. 1866. General Morphologie der Organismen. Berlin: G. Reimer.

Haeckel E. 1905. Der Kampf um den Entwickelungs-Gedanken. Berlin: G. Reimer.

Hennig W. 1966. Phylogenetic systematics. Translated by D. Dwight Davis and Rainer Zangerl. Urbana: Univ Illinois Press.

Hunn E. 1975. A measure of the degree of correspondence of folk to scientific biological classification. American Ethnologist 2:309–327.

Huxley J. 1957. New bottles for new wine: Essays. London: Chatto & Windus.

Huxley TH. 1872. Anatomy of vertebrated animals. New York: Appleton.

Illiger JKW. 1811. Prodomus systematis mammalium et avium additis terminis zoographicii utriusque classis eorumque versione germanica. Berlin: C. Salfeld.

Kirsch JA, Dickerman AW, Reig OA, Springer MS. 1991. DNA hybridization evidence for the Australasian affinity of the American marsupial Dromiciops australis. Proc Natl Acad Sci USA 88: 10465–10469.

Kjer KM, Honeycutt RL. 2007. Site specific rates of mitochondrial genomes and the phylogeny of eutheria. BMC Evol Biol 7: 8.

Klein JT. 1751. Quadrupedum dispositio brevisque historia naturalis. Leiden: BC Bretikopf.

Lacepede B. 1799. Tableau des divisions, sousdivisions, ordres et genres des Mam mifères, des Cétacés et des Oiseaux. Paris: Plassan.

Linnaeus C von. 1766. Systema Naturae. Editio Duodecima, Reformata. Stockholm: Laurentii Salvii.

Maddison WP, Maddison DR. 2015. Mesquite: a modular system for evolutionary analysis. Software published by the authors, http://mesquiteproject.wikispaces.com/

McCormack JE, Faircloth BC, Crawford NG, Gowaty PA, Brumfield RT, Glenn TC. 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res 22: 746–754.

McKenna MC. 1975. Toward a phylogenetic classification of the Mammalia. In Phylogeny of the Primates (eds. W.P. Luckett and F.S. Szalay), pp. 21–46, Springer.

McKenna MC. 1991. The alpha crystallin A chain of the eye lens and mammalian phylogeny. Annales Zoologici Fennici 28:349–360.

McKenna MC, Bell SK. 1997. Classification of Mammals Above the Species Level. New York: Columbia Univ Press.

Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TLL, Stadler T, et al. 2011. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. Science 334: 521–4.

Mitchell KJ, Pratt RC, Watson LN, Gibb GC, Llamas B, Kasper M, Edson J, Hopwood B, Male D, Armstrong KN, et al. 2014. Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. Mol Biol Evol 31: 2322–2330.

Miyamoto MM, Goodman M. 1986. Biomolecular systematics of eutherian mammals: phylogenetic patterns and classification. Syst Biol 35: 230–240.

Montgomery JE. 2013. Al-Jahiz: In Praise of Books. Edinburgh: Edinburgh University Press.

Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder O a, Stanhope MJ, de Jong WW, Springer MS. 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294: 2348–2351.

Novacek MJ. 1992. Mammalian phylogeny: shaking the tree. Nature 356: 121–125.

Novacek MJ. 1986. The skull of leptictid insectivorans and the higher-level classification of eutherian mammals. Bulletin of the American Museum of Natural History 183: 1–112.

Numbers RL. 2006. The Creationists, Expanded Edition. Cambridge MA: Harvard University Press.

Osborn HF. 1917. The Origin and Evolution of Life. New York: Scribners.

Owen R. 1868. On the Anatomy of Vertebrates, vol. 3 Mammals. London: Longman Green.

Patterson C. 1988. The impact of evolutionary theories on systematics. In Prospects in systematics (ed. D.L. Hawksworth), pp. 59–91, Clarendon Press, Oxford.

Pennant T. 1781. History of quadrupeds. London: B. White.

Penny D, Foulds LR, Hendy MD. 1982. Testing the theory of evolution by comparing phylogenetic trees constructed from five different protein sequences. Nature 297: 197–200.

Pietsch TW. 2012. Trees of Life: a visual history of evolution. Baltimore: Johns Hopkins Univ Press.

Prothero DR. 2007. Evolution: what the fossils say and why it matters. New York: Columbia University Press.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.

Ray J. 1693. Synopsis methodica animalium quadrupedum et serpentini generis. London: Smith & Walford.

Rieppel O. 1986. Species are individuals. Evolutionary Biology 20: 283–317.

Romer AS. 1959. The Vetebrate Story. Chicago: Univ Chicago Press.

Romer AS. 1945. Vertebrate Paleontology. Chicago: Univ Chicago Press.

Rudwick MJS. 1997. Fossil Bones & Geological Catastrophes. Chicago: Univ Chicago Press.

Scopoli GA. 1777. Introductio ad historiam naturalem: sistens genera lapidum, plantarum, et animalium hactenus detecta, caracteribus essentialibus donata in tribus divisa, subinde ad leges naturae. Prague: Wolfgang Gerle.

Shoshani J, McKenna MC. 1998. Higher taxonomic relationships among extant mammals based on morphology, with selected comparisons of results from molecular data. Mol Phylogenet Evol 9: 572–84.

Simpson GG. 1931. A new classification of mammals. Bulletin of the American Museum of Natural History. 59: 259-293.

Simpson GG. 1944. Tempo and mode in evolution. New York: Columbia Univ Press.

Simpson GG. 1945. The principles of classification and a classification of mammals. Bulletin of the American Museum of Natural History 85:1-350.

Sober E. 2008. Evidence and Evolution: the Logic Behind the Science. Cambridge: Cambridge University Press.

Song S, Liu L, Edwards S V, Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proc Natl Acad Sci USA 109: 14942–14947.

Stanhope MJ, Waddell VG, Madsen O, de Jong W, Hedges SB, Cleven GC, Kao D, Springer MS. 1998. Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. Proc Natl Acad Sci USA 95: 9967–9972.

Storr GCC. 1780. Prodromus methodi mammalium... inauguralem disputationem. Tübingen: F Wolffer.

Stott R. 2012. Darwin's Ghosts: the Secret History of Evolution. London: Bloomsbury.

Tarver JE, dos Reis M, Mirarab S, Moran RJ, Parker S, O'Reilly JE, King BL, O'Connell MJ, Asher RJ, Warnow T, Peterson KJ, Donoghue PCJ, Pisani D. 2016. The Interrelationships of Placental Mammals and the Limits of Phylogenetic Inference. Genome Biol Evol 8(2):330–344.

Teeling EC, Springer MS, Madsen O, Bates P, O'Brien SJ, Murphy WJ. 2005. A molecular phylogeny for bats illuminates biogeography and the fossil record. Science 307: 580–584.

Van Valen L. 1978. Why not to be a cladist. Evol Theory 3: 285–299.

Wagner JA. 1855. Schreber's Die Säugethiere, suppl. Vol. 5. Leipzig: Weigel.

Weber MWC. 1904. Die Säugetiere: Einführung in die Anatomie und Systematik der Recenten und Fossilen Mammalia. Jena: Gustav Fischer.

Winge H. 1941. The Interrelationships of the Mammalian Genera. Copenhagen: CA Reitzels.

Wray GA, Hoekstra HE, Futuyma DJ, Lenski RE, Mackay TFC, Schluter D, Strassmann JE. 2014. Does evolutionary theory need a rethink? Nature 514: 161–164.

Tables

Table 1. Sources of branching diagrams used to represent historical classifications.

Author	**Year**	**Publication**	**Source**	**Method**
Leviticus 11	~ -500	*The Bible*	Gregory 1910:7-8	ancient
Al-Jahiz	~ 850	*Kitab al Hayawan*	Eisenstein 1991	ancient
Ray	1693	*Synopsis Methodica Animalium...*	Gregory 1910:18-19	pre-evolutionary
Linnaeus	1735	*Systema Naturae I ed.*	Gregory 1910:102	pre-evolutionary
Klein	1751	*Quadrupedum Dispositio Brevisque Historia Naturalis*	Gregory 1910:26-27	pre-evolutionary
Brisson	1762	*Regnum Animale... II ed.*	Gregory 1910:42-43	pre-evolutionary
Linnaeus	1766	*Systema Naturae XII ed.*	Gregory 1910:29-30	pre-evolutionary
Scopoli	1777	*Introductio ad Historium*	Gregory 1910:37	pre-evolutionary
Blumenbach	1779	*Handbuch der Naturgeschichte*	Gregory 1910:45-46	pre-evolutionary
Storr	1780	*Prodromus Methodi Mammalium*	Gregory 1910:49-50	pre-evolutionary
Pennant	1781	*History of Quadrupeds*	Gregory 1910:51-52	pre-evolutionary
Lacépède	1799	*Tableu des Divisions...*	Gregory 1910:62-63	pre-evolutionary
Cuvier	1800	*Leçons d'Anatomie Comparée*	Gregory 1910:65-66	pre-evolutionary
Illiger	1811	*Prodromus Mammalium...*	Gregory 1910:70-71	pre-evolutionary
DeBlainville	1816	Prodrome d'une nouvelle distribution...	Gregory 1910:77-78	pre-evolutionary
Cuvier	1817	*Le Règne Animal*	Gregory 1910:80	pre-evolutionary
Blumenbach	1830	*Handbuch der Naturgeschichte 12. ed*	Gregory 1910:81	pre-evolutionary
Deblainville	1834	Gervais, *Mammalogie ou Mastologie...*	Gregory 1910:82-83 and Guerin 1836:619	pre-evolutionary
Bonaparte	1837	*A new systematic arrangement...*	Gregory 1910:84 and Bonaparte 1840	pre-evolutionary
Wagner	1855	*Schreber's Säugetiere*	Gregory 1910:86-87	pre-evolutionary
Haeckel	1866	*Anthropogenie...*	Pietsch 2012: 119	evolutionary

Author	Year	Publication	Source	Method
Owen	1868	*On the Anatomy of Vertebrates*	Gregory 1910:90	pre-evolutionary
Gill	1870	*On the Relations of the Orders of Mammals*	Gregory 1910:92-93	evolutionary
Huxley	1872	*A Manual of the Anatomy...*	Gregory 1910:93	evolutionary
Flower	1883	*On the Arrangement...*	Gregory 1910:96-97	evolutionary
Cope	1898	*Syllabus of Lectures on Geology and Paleontology...*	Gregory 1910:98-99	evolutionary
Gadow	1898	*Classification of the Vertebrata...*	Gadow 1898:39-54	evolutionary
Weber	1904	*Die Säugetiere*	Gregory 1910:100-101	evolutionary
Haeckel	1905	*Der Kampf um den Entwicklungsgedanken*	Pietsch 2012: 121	evolutionary
Gregory	1910	*The Orders of Mammals*	Gregory 1910:466-467	evolutionary
Osborn	1917	*Origin and Evolution of Life*	Pietsch 2012:169	evolutionary
Winge	1921*	*The Interrelationships of the Mammalian Genera*	Winge 1941	evolutionary
Cabrera	1922	*Manual de Mastozoologia*	Cabrera 1922	evolutionary
Simpson	1931	*A New Classification of Mammals*	Simpson 1931	evolutionary
Romer	1945	*Vertebrate Paleontology,* 1st (1945) and 3rd (1966) editions	Pietsch 2012:190-194	evolutionary
Simpson	1945	*Principles of Classification...*	Simpson 1945	evolutionary
Grassé	1955	*Traité de Zoologie vol. 17*	Piveteau 1955:8	evolutionary
Romer	1959	*The Vertebrate Story*	Romer 1959	evolutionary
Mckenna	1975	chapter 2 in *Phylogeny of the Primates* (Luckett & Szalay)	McKenna 1975	cladistic
Miyamoto	1986	*Syst Zool* 35(2):230-240	Miyamoto & Goodman 1986: fig. 3	molecular
Novacek	1986	*The Skull of Leptictids...*	Novacek 1986: fig. 35	cladistic
Novacek	1992	*Nature* 356:121-125	Novacek 1992: fig. 1	cladistic

Author	Year	Publication	Source	Method
Mckenna	1997	*Classification of Mammals*	McKenna & Bell 1997	cladistic
Shoshani	1998	*Mol Phy Evol* 9(3):572-584	Shoshani & McKenna 1998: fig. 1	cladistic
Stanhope	1998	*Proc Nat Acad Sci USA* 95:9967-9972	Stanhope et al. 1998: fig. 1	molecular
Murphy	2001	*Science* 294:2348-2351	Murphy et al.2001:fig.1	molecular
Arnason	2002	*Proc Nat Acad Sci USA* 99:8151-8156	Arnason et al. 2002: fig. 1	molecular
Asher	2003	*J Mamm Evol* 10:131-194	Asher et al. 2003: fig. 5 (left)	combined
Kjer	2007	*BMC Evol Biol* 7:8	Kjer & Honeycutt 2007: fig. 1	molecular
Prothero	2007	*Evolution: What the Fossils Say and Why it Matters*	Prothero 2007:285,291, 298, 309, 338, 312, 324	cladistic
Arnason	2008	*Gene* 412:37-51	Arnason et al. 2008: fig. 2	molecular
Meredith	2011	*Science* 334:521-524	Meredith et al. 2011: fig. 1	defines correct tree
Mccormack	2012	*Genome Research* 22:746-754	McCormack et al. 2012: fig. 2	molecular
Song	2012	*Proc Nat Acad Sci USA* 109:14942-47	Song et al. 2012: fig. 1	molecular
Tarver	2016	*Genome Biol Evol* 8:330-344	Tarver et al. 2016:fig. 2	defines correct tree

* As noted in the text, Winge's taxonomic views were forumlated at latest during the early 1920s, not the (posthumous) 1941 date of publication.

Table 2. Translation of Ray 1693: 53, 60-61 (original in Fig. 2).

General Table of Animals

Animals are either
- with blood
  - and breathing with lungs and possessing in the heart
    - two ventricles
      - and live birth
        - and are aquatic (Cetacea)
        - or terrestrial and quadrupedal, and also possess hair, as in ["Manati"]. Animals whose genera we count as terrestrial amphibians.
      - or are egg-laying (Aves)
    - one ventricle of the heart: quadrupeds that lay eggs* and serpents
  - gills for breathing: also all marine fish with blood
- without blood

Table of quadrupedal animals with live birth

Hairy animals with live birth or four legs are either

Ungulates, those with either
- one-toed: horse, ass, zebra
- double-toed: bisulca, or split hoof, or those which
  - Ruminantia with horns
    - perpetual: cows, sheep, goats
    - periodic: deer
  - non-ruminant, swine
- four-toed: rhino, hippo, etc.

Unguiculata, whose feet are either
- only two toes: camel
- many toes: which are either
  - undivided digits attached to a common roof of skin with the margins of the foot bearing weight, strong and blunt digits: elephant
  - when divided one from the other are either
    - flat-nailed and anthropomorphous: simians
    - narrow-nailed, teeth or incisors in each maxilla, either
      - several or all carnivorous and predatory, at least insectivora or common diet of insects and plants
        - greater rostrum
          - short, round head: cats
          - long: dogs
        - lesser, long slender body, short legs: weasels
      - two important kinds, whose diet is herbivorous: hares

Among the live-bearing, multi-toed, anomalous quadrupeds are the hedgehog, tatou or armadillo, mole, mouse, anteater, bat, and sloth. The former five have a lengthened rostrum like the dog, but the verminous species have teeth of different form and disposition. Indeed the anteater completely lacks teeth. In the latter two the rostrum is short.

**Gregory (1910:18) notes that Ray's original text states "viviparous" (see Fig. 2) but points out that this is a lapsus and that Ray's intent is to identify egg-laying animals.*

Table 3. Median similarity (using "partitions in common" in Mesquite [Maddison & Maddison 2015]) among cladograms to each other, derived from classifications within 12 discrete time bins.

time	dominant method	# studies	median # of partitions in common with each other	median actual/possible partitions in common with each other
1751 & before	pre-evolutionary	5	0.0	0.000
1762-1780	pre-evolutionary	5	1.0	0.033
1781-1816	pre-evolutionary	5	3.0	0.085
1817-1855	pre-evolutionary	5	5.0	0.129
1866-1883	evolutionary	5	13.0	0.343
1898-1910	evolutionary	5	12.5	0.276
1917-1931	evolutionary	4	15.5	0.286
1945-1975	evolutionary	5	23.0	0.411
1986-1997	cladistic	4	17.5	0.417
1998-2002	molecular	4	12.5	0.279
2003-2008	molecular	4	19.0	0.461
2011-2016	molecular	4	16.0	0.572

Table 4. Significance values of a pairwise Wilcoxon rank sum test across mean similarities of cladograms derived from classifications based on different methodologies (with sample size over two) to the well-corroborated tree. Significant differences are shown in bold. ("pairwise.wilcox.test(treegoodsample$actPot,treegoodsample$method,p.adj='bonferroni')")

	Molecular	Cladistic	Evolutionary
Pre-evolutionary	**0.00036**	**0.00191**	**0.00003**
Evolutionary	**0.00159**	0.06846	-
Cladistic	0.10010	-	

Figure Captions

Figure 1. Above: Well-corroborated tree of mammalian relationships based on congruence between optimal topologies from Meredith et al. (2011), Fabre et al. (2012), Mitchell et al. (2014), and Tarver et al. (2016). Below: branching diagram based on classification of Storr (1780); groups in agreement with the well-corroborated tree have thick branches. Abbreviations (above) represent high-level taxa to the right of the respective node (nomenclature following Asher & Helgen 2010): Af = Afrotheria, Afi = Afroinsectiphilia, Ar = Artiodactyla, At = Atlantogenata, Au = Australodelphia, Ca = Carnivora, Ch = Chiroptera, Eu = Eucarchontoglires, Fe = Ferae, Gl = Glires, Ha = Haplorhini, La = Laurasiatheria, Li = Lipotyphla, Ma = Marsupialia, Mo = Monotremata, Pa = Paenungulata, Pe = Perrisodactyla, Pi = Pilosa, Pl = Placentalia, Pr = Primates, Ro = Rodentia, Ru = Ruminantia, Th = Theria, Xe = Xenarthra. Selected groups are color coded (red = Euarchontoglires, blue = Laurasiatheria, green = Atlantogenata, yellow = monotremes & marsupials) for ease of reference to subsequent figures. Dark red indicates "Archonta" (i.e., Primates, Dermoptera, Scandentia), the monophyly of which is uncertain (as indicated in the tree).

Figure 2. A: Reproduction of Ray (1693) pp. 60-61 ("Animalium Viviparorum Quadrupedum Tabula") and p. 53 ("Animalia Tabula Generalis"), reprinted in Gregory (1910:18-19) and translated in Table 2.

Figure 3. Ratio of actual by potential number of groups (Y-axis) held in common with the well-corroborated tree (Fig. 1) by year (X-axis), distinguishing data points and trend lines for different methods (asterisk = ancient, black circle = pre-evolutionary, red triangle = evolutionary, green plus = cladistic, aqua diamond = combined, blue x = molecular). Selected authors are indicated. Sample sizes (n) and Pearson's correlation coefficients are given for each category in parentheses at top left; parametric significance of the correlation below 0.05 is indicated with an asterisk. Note the discontinuity in the X-axis prior to 1693.

Figure 4. Agreement among authors within 12 time bins, calculated as the median value of each pairwise comparison of actual / potential number of groups in common with each other (Table 3). Higher values on the Y-axis represent greater levels of inter-author agreement.

Figure 5. Box plots showing average (thick horizontal line), middle quartiles (box) and range (dashed vertical lines) of ratio of actual / potential number of groups held in common with the well-

corroborated tree (Y-axis) across five methods. Higher values on the Y-axis represent greater levels of agreement with the well-corroborated tree (Fig. 1).

Figure 6. Cladograms extracted from classifications by A) Cuvier 1817, B) DeBlainville 1834 (see also Fig. 7), C) Gregory 1910, and D) McKenna & Bell 1997. Selected groups (as defined in Fig. 1) are color coded (red = Euarchontoglires, blue = Laurasiatheria, green = Atlantogenata, yellow = monotremes & marsupials). Groups present in the well-corroborated tree (Fig. 1) are shown with thick branches.

Figure 7. Facsimilie of DeBlainville's 1834 classification, from p. 619 of the *Dictionnaire Pittoresque d'Histoire Naturelle* (Guérin 1836).

Appendix 1. Treefile in nexus format with topologies derived from all 55 studies listed in Table 1: http://people.ds.cam.ac.uk/rja58/trees-appen1.nex